



The Standardization Survival Kit (SSK)

Marie Puren, Charles Riondet, Laurent Romary, Dorian Seillier, Lionel Tadjou

► To cite this version:

Marie Puren, Charles Riondet, Laurent Romary, Dorian Seillier, Lionel Tadjou. The Standardization Survival Kit (SSK): Bringing best practices to research communities in the Humanities. Digital Humanities Benelux 2018, Jun 2018, Amsterdam, Netherlands. hal-01850075

HAL Id: hal-01850075

<https://hal.science/hal-01850075>

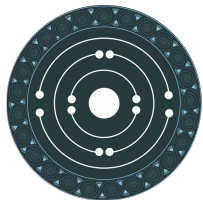
Submitted on 26 Jul 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution| 4.0 International License



PARTHENOS

Pooling Activities, Resources and Tools
for Heritage E-research Networking,
Optimization and Synergies



DH Benelux 2018, 6–8 June 2018

The Standardization Survival Kit (SSK)

Bringing best practices to research communities in the
Humanities



Marie Puren, Charles Riondet, Laurent Romary, Dorian Seillier, Lionel Tadjou
Inria, PARTHENOS



Why using **standards**?

Manage your data, don't let your data manage you...



... and produce **interoperable**, **sustainable** & **reusable** results for research.

Why using **standards**?

*“Standards are a **key to great digital research**, which helps to discover and understand our cultural and societal life.”*

[Laurent Romary](#)

Chairman of the Technical Committee "Terminology and other language and content resources" of the International Organization for Standardization (ISO)



What are **standards**?

To be called "standards", they must fulfill three requirements:

- They express a **consensus**
- They are published and easily **accessible**
- They are **maintained**



The SSK : a toolkit for Humanities scholars

<https://ssk-application.parthenos.d4science.org/ssk/#/>



PARTHENOS
PARIS ARCHAEOLOGICAL RESEARCH
THROUGH OPEN SCIENCE

HOME

EXPLORE

CONTRIBUTE

PARTNERS

GLOSSARY

MAIN SITE

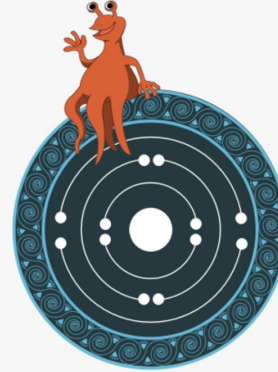
Log in



SSK/ >

The Standardization Survival Kit :
for a wider use of standards within Arts and
Humanities

BROWSE BY ▾



WHY STANDARDS ?



Standards are generally published by standardization organizations, such as ISO, W3C or the TEI Consortium.



They express a consensus.



They are accessible to anyone.



The SSK : a toolkit for **Humanities** scholars

- Documenting



- Supporting



- Training



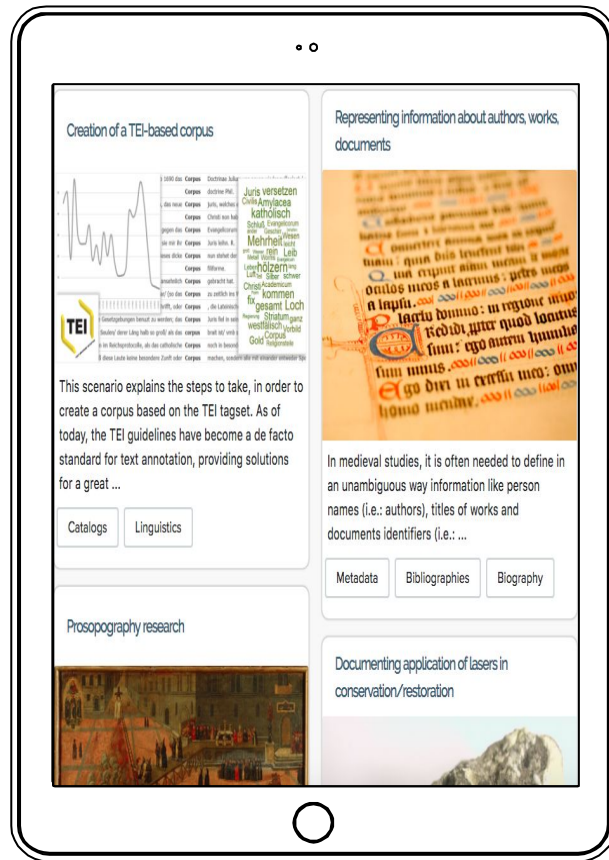
- Communicating





The scenarios

Providing **contextual information** and **relevant examples** on how standards can be applied in a given research project.



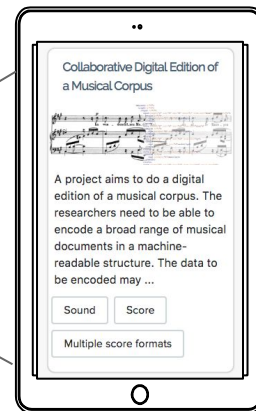


What Tork & Mork tell us

Hey Mork ! With some of my **colleagues**, I'd like to put **musical scores online**. But I don't know **what I can do** and how to proceed! Can you help me?



Well Tork, **use the SSK!** I think the scenario "**Collaborative Digital Edition of a Musical Corpus**" can be very useful for you. It will guide you through all the **necessary steps** to put these scores online, and present the **best practices** associated to each of these steps.





A high-level research guide

Collaborative Digital Edition of a Musical Corpus

A project aims to do a digital edition of a musical corpus. The researchers need to be able to encode a broad range of musical documents in a machine-readable structure. The data to be encoded may include the musical content as provided by the composer (notes, pitches, durations, dynamics, etc.), information on the score (incipit, lyrics writer, etc.), information added by a performer when interpreting the content (timing, phrasing, various annotations, etc.), information on the visual appearance of the score (page layout, musical font, etc.) and analyses of the content in any of the other domains. The edition will be structured around a database in order to allow the users to explore it more easily. Furthermore, the project intends to be collaborative, which means it will offer anyone interested the possibility to contribute.



Techniques ?

Encoding



1 . Create a digital corpus of musical compositions.

Select resources to be included in the corpus. After collecting original musical sources, transcribe them adding critical editorial signs and normalizing, where applicable, ancient poetic texts to modern usage. To get directly MEI files, use MEISE (MEI Score Editor).



Resources and best practices

GENERAL RESOURCES



tool
Avid Technology
Sibelius



documentation
Perry Roland, Johannes Kepper
Music Encoding Initiative Guidelines
"2016"



schema
MEI Schemata



documentation
MEI Score Editor - MEISE



documentation
MEISEnNative note editor for the MEI encoding format



tutorial
MEISE Wiki



PROJECT-SPECIFIC RESOURCES



paper
Laurent Romary, Charles Riondet
EAD-ODD: A solution for project-specific EAD schemes
"2018/03/21"



paper
Laurent Romary, Charles Riondet
Towards multiscale archival digital data
"2017/09/12"



tutorial
XPath and Schematron for TEI Customization



report
hal



code

Three layers



Scenario

A complete and generic research use case composed of several steps to be followed.



Step

A unique task to be performed inside a scenario with the help and recommendation of one or several resources.



Resource

A standardized tool / service / document guiding the researcher in her/his tasks completion.

Managing resources

Standards documentation
(ISO, TEI)

Official publications and
reports (D4Science, HAL,
Zenodo)

Documentation



Reference libraries
organised by domains and
standards

Maintenance with Zotero

Bibliography



Code snippets (GitHub)

Tools & services
(D4Science)

Technical
resources



Wikis

Blog posts
(Hypotheses.org)

Discussion lists

User
communities



The TEI model

The scenarios are described using the TEI format (Text Encoding Initiative). All the information displayed within the SSK proceed from TEI files.

```
<text>
<body>
  <div type="scenario" xml:id="sc_schemaCustomization">
    <head xml:lang="en" type="scenarioTitle">Harmonization of digitized textual resources
      with the DTABf</head>
    <desc type="definition" xml:lang="en">Currently, initiatives for the digitization of
      textual resources and their provision to the interested community are manifold and
      various. Hence, scholars who want to base their research on digitized texts,
```

```
<listEvent>
  <event ref="step_OtICoP_171117" type="researchStep" xml:id="s1"/>
  <event ref="step_CaC_171117" type="researchStep" xml:id="s2"/>
  <event ref="step_SaD_171117" type="researchStep" xml:id="s3"/>
  <event ref="step_A_171117" type="researchStep" xml:id="s4"/>
  <event ref="step_CiSF_171117" type="researchStep" xml:id="s5"/>
  <event ref="step_TtI_171117" type="researchStep" xml:id="s6"/>
```

What is next?

- **Browsing** vocabularies and bibliography
- Creating an **account** :
 - to manage **bookmarks**
 - to customize **scenarios** (by combining existing steps from SSK's research scenarios)
- **Contributing** (in TEI or directly on the interface) :
 - creation
 - edition
 - customization
- Accessing a **multilingual** interface



Test the SSK!



<https://ssk-application.parthenos.d4science.org/ssk/>



ssk@inria.fr



Why standards?: <https://youtu.be/ULNztgzTKkU>

Thank you for your attention!

 <http://www.parthenos-project.eu/>

 @Parthenos_EU

 @PARTHENOSproject

 contact@parthenos-project.eu